

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.**

Application Serial Number: 10/762,049  
Source: 1Fw16  
Date Processed by STIC: 8/22/05

# ***ENTERED***



IFW16

## RAW SEQUENCE LISTING

DATE: 08/22/2005

PATENT APPLICATION: US/10/762,049

TIME: 12:16:16

Input Set : N:\Crf3\RULE60\10762049.raw

Output Set: N:\CRF4\08222005\J762049.raw

1 <110> APPLICANT: Stephen M. Allen  
 2 Saverio C. Falco  
 3 Catherine J. Thorpe  
 4 <120> TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
 5 <130> FILE REFERENCE: BB-1167  
 6 <140> CURRENT APPLICATION NUMBER: 10/762,049  
 7 <141> CURRENT FILING DATE: 2004-01-21  
 8 <150> PRIOR APPLICATION NUMBER: US/09/720,317  
 9 <151> PRIOR FILING DATE: 2000-12-21  
 10 <150> PRIOR APPLICATION NUMBER: 60/092,833  
 11 <151> PRIOR FILING DATE: 1998-07-14  
 12 <160> NUMBER OF SEQ ID NOS: 31  
 13 <170> SOFTWARE: Microsoft Office 97  
 15 <210> SEQ ID NO: 1  
 16 <211> LENGTH: 2279  
 17 <212> TYPE: DNA  
 18 <213> ORGANISM: Zea mays  
 19 <400> SEQUENCE: 1

fb

20	gcacgagaga	agaagaagga	agggccatct	tccgaccac	ttgtaggcgc	tgtaagcctg	60
21	taagcagtca	cctctcagcc	acagacatgg	tgggcatgag	aggcgccctac	ggtggtgctt	120
22	gcaatgacga	cagcaagagc	cggctgcacg	gaggcaaggc	ggcggagccg	gagatcgctg	180
23	cgatggcagt	gcacaaggtg	gcgcccac	cggcgccgag	cacggcgagc	aagatgaagg	240
24	tgagggtgaa	ggagaccttc	ttccccgacg	acccgttccg	ggcggtcaag	gggcagccgc	300
25	cggggacgca	gtgggtcatg	gcgggtcaggt	acctcttccc	catcctggac	tgggtgccga	360
26	gctactcctt	gtcgtctctt	aagtccgacc	tcgctcgagg	cctcaccatt	gccagcctcg	420
27	ccattcctca	gggcattagc	tacgcgaagc	tggcaagctt	gcctcccata	atcggtgctg	480
28	attcgagctt	cgtgccgcgc	atggtgtacg	cgggtgctgg	gagctcccgt	gacctggcgg	540
29	tgggcccggg	gtcgatctcg	tcgctgatca	tgggggtccat	gctgcggcag	gccgtgagcc	600
30	ccactgcgga	gccgacgctg	ttcctgcagc	tggccttcac	ctccaccctg	ttcgcggggc	660
31	tgggtgcaggc	ctccctgggc	atcctcaggc	tcggcttcgt	catcgacttc	ctgtccaagg	720
32	cgacgtggtg	gggggttcag	gccggcgccg	ccatcatcgt	ggcgctgcag	caactcaagg	780
33	ggctgctggg	catcgtccac	ttcaccaccg	agatgggcat	cgtcccagtc	atggcctccg	840
34	tcttccacca	caccagcgag	tgggtcgtggc	agacgatact	catgggctgc	tgcttctctg	900
35	tcttctctgt	gtcggcgagg	catgtgagca	tcagatggcc	aaagcttttc	tgggtttcgg	960
36	cgtgcgcgcc	cctggcatcg	gtcaccatct	cgacgtgctt	tgttttctct	ttcaaagctc	1020
37	agaaccatgg	catcagcatc	attgggcagc	tcaagtgcgg	cctgaatcgc	ccctcgtggg	1080
38	acaagctcct	gtttgacacg	gcgtatttag	gcctcaccat	gaagactggc	cttgtcaccg	1140
39	gaatcatctc	actgacggaa	ggaatagcgg	ttggtagaac	atttgctca	ctcaaggact	1200
40	accagataga	tggaaacaag	gagatgatgg	ccatagggtt	gatgaatgtt	gttgggtcct	1260
41	gcacatcatg	ctacgtaaca	acaggtgcgt	tctcccgcct	tgctgtaaac	cacaacgccg	1320
42	gctgcaagac	tgccatgtcc	aacgtgatca	tggcgctgac	tgtgatggtc	acgctgctgt	1380
43	tcctcatgcc	actgttcgtg	tacacacca	acgttgtcct	cggagcgatc	atcatcgccg	1440
44	cgggtgatcgg	cctgatcgat	ttccccgcgg	tgtaccacat	ctggaagatg	gacaagatgg	1500

## RAW SEQUENCE LISTING

DATE: 08/22/2005

PATENT APPLICATION: US/10/762,049

TIME: 12:16:16

Input Set : N:\Crif3\RULE60\10762049.raw

Output Set: N:\CRF4\08222005\J762049.raw

```

45 attttctggt gtgcgtttgc gcgtttgccg gcgtcatctt catctcagtc caagaaggcc 1560
46 ttgcgatagc ggttggtata tctatattta ggggtgtgat gcagatcaca aggccgaaga 1620
47 tgatggttca agggaacatc aaggggactg atatttacag agacctgcat cactacaagg 1680
48 aggcccaaag agtttctggg ttcttgatct tggccattga agcaccgata aacttcgcca 1740
49 actccaacta cctgaatgaa aggattaaaa gatggataga ggaagaatct tttgaacagg 1800
50 ataaacatac tgaactccat ttcataatct tggatctgtc agctgttcct gcaattgaca 1860
51 caagtggcat agcgttcctc attgacataa agaaatcaat agagaaacgt ggtctggagc 1920
52 ttgtgcttgt caatccaact ggagaagtca tggagaaaat acaacgtgca aacgaggctg 1980
53 aaaactattt taggccagat tgcttgatc tgaccactgg cgaagcaatc gcttcacttt 2040
54 ctgcacttgc caagatgaca aaaccctaaa tggattgctg aattgtcatt gtgttcattc 2100
55 ctagcactgt taaaagtttt cgggtgcagga ttttctgtaa tggggagtgc atccaatagg 2160
56 agtacatcac agctatgttt gtatctagta gaattcttca gatccatgtg atgcaaattc 2220
57 aatggaaaac aaatatgaca gtacaatagt agatcttaca gaaattttct gctgcaaaa 2279

```

59 &lt;210&gt; SEQ ID NO: 2

60 &lt;211&gt; LENGTH: 688

61 &lt;212&gt; TYPE: PRT

62 &lt;213&gt; ORGANISM: Zea mays

63 &lt;400&gt; SEQUENCE: 2

```

64 Thr Arg Glu Glu Glu Gly Arg Ala Ile Phe Arg Pro Thr Cys Arg Arg
65 1 5 10 15
66 Cys Lys Pro Val Ser Ser His Leu Ser Ala Thr Asp Met Val Gly Met
67 20 25 30
68 Arg Gly Ala Tyr Gly Gly Ala Cys Asn Asp Asp Ser Lys Ser Arg Leu
69 35 40 45
70 His Gly Gly Lys Ala Ala Glu Pro Glu Ile Ala Ser Met Ala Val His
71 50 55 60
72 Lys Val Ala Pro Pro Pro Ala Arg Ser Thr Ala Ser Lys Met Lys Val
73 65 70 75 80
74 Arg Val Lys Glu Thr Phe Phe Pro Asp Asp Pro Phe Arg Ala Phe Lys
75 85 90 95
76 Gly Gln Pro Pro Gly Thr Gln Trp Leu Met Ala Val Arg Tyr Leu Phe
77 100 105 110
78 Pro Ile Leu Asp Trp Val Pro Ser Tyr Ser Leu Ser Leu Phe Lys Ser
79 115 120 125
80 Asp Leu Val Ala Gly Leu Thr Ile Ala Ser Leu Ala Ile Pro Gln Gly
81 130 135 140
82 Ile Ser Tyr Ala Lys Leu Ala Ser Leu Pro Pro Ile Ile Gly Leu Tyr
83 145 150 155 160
84 Ser Ser Phe Val Pro Pro Met Val Tyr Ala Val Leu Gly Ser Ser Arg
85 165 170 175
86 Asp Leu Ala Val Gly Pro Val Ser Ile Ser Ser Leu Ile Met Gly Ser
87 180 185 190
88 Met Leu Arg Gln Ala Val Ser Pro Thr Ala Glu Pro Thr Leu Phe Leu
89 195 200 205
90 Gln Leu Ala Phe Thr Ser Thr Leu Phe Ala Gly Leu Val Gln Ala Ser
91 210 215 220
92 Leu Gly Ile Leu Arg Leu Gly Phe Val Ile Asp Phe Leu Ser Lys Ala
93 225 230 235 240
94 Thr Leu Val Gly Phe Met Ala Gly Ala Ala Ile Ile Val Ala Leu Gln

```

## RAW SEQUENCE LISTING

DATE: 08/22/2005

PATENT APPLICATION: US/10/762,049

TIME: 12:16:16

Input Set : N:\CrF3\RULE60\10762049.raw

Output Set: N:\CRF4\08222005\J762049.raw

```

95          245          250          255
96  Gln Leu Lys Gly Leu Leu Gly Ile Val His Phe Thr Thr Glu Met Gly
97          260          265          270
98  Ile Val Pro Val Met Ala Ser Val Phe His His Thr Ser Glu Trp Ser
99          275          280          285
100  Trp Gln Thr Ile Leu Met Gly Val Cys Phe Leu Val Phe Leu Leu Ser
101          290          295          300
102  Ala Arg His Val Ser Ile Arg Trp Pro Lys Leu Phe Trp Val Ser Ala
103          305          310          315          320
104  Cys Ala Pro Leu Ala Ser Val Thr Ile Ser Thr Leu Leu Val Phe Leu
105          325          330          335
106  Phe Lys Ala Gln Asn His Gly Ile Ser Ile Ile Gly Gln Leu Lys Cys
107          340          345          350
108  Gly Leu Asn Arg Pro Ser Trp Asp Lys Leu Leu Phe Asp Thr Ala Tyr
109          355          360          365
110  Leu Gly Leu Thr Met Lys Thr Gly Leu Val Thr Gly Ile Ile Ser Leu
111          370          375          380
112  Thr Glu Gly Ile Ala Val Gly Arg Thr Phe Ala Ser Leu Lys Asp Tyr
113          385          390          395          400
114  Gln Ile Asp Gly Asn Lys Glu Met Met Ala Ile Gly Leu Met Asn Val
115          405          410          415
116  Val Gly Ser Cys Thr Ser Cys Tyr Val Thr Thr Gly Ala Phe Ser Arg
117          420          425          430
118  Ser Ala Val Asn His Asn Ala Gly Cys Lys Thr Ala Met Ser Asn Val
119          435          440          445
120  Ile Met Ala Leu Thr Val Met Val Thr Leu Leu Phe Leu Met Pro Leu
121          450          455          460
122  Phe Val Tyr Thr Pro Asn Val Val Leu Gly Ala Ile Ile Ile Ala Ala
123          465          470          475          480
124  Val Ile Gly Leu Ile Asp Phe Pro Ala Val Tyr His Ile Trp Lys Met
125          485          490          495
126  Asp Lys Met Asp Phe Leu Val Cys Val Cys Ala Phe Ala Gly Val Ile
127          500          505          510
128  Phe Ile Ser Val Gln Glu Gly Leu Ala Ile Ala Val Gly Ile Ser Ile
129          515          520          525
130  Phe Arg Val Leu Met Gln Ile Thr Arg Pro Lys Met Met Val Gln Gly
131          530          535          540
132  Asn Ile Lys Gly Thr Asp Ile Tyr Arg Asp Leu His His Tyr Lys Glu
133          545          550          555          560
134  Ala Gln Arg Val Ser Gly Phe Leu Ile Leu Ala Ile Glu Ala Pro Ile
135          565          570          575
136  Asn Phe Ala Asn Ser Asn Tyr Leu Asn Glu Arg Ile Lys Arg Trp Ile
137          580          585          590
138  Glu Glu Glu Ser Phe Glu Gln Asp Lys His Thr Glu Leu His Phe Ile
139          595          600          605
140  Ile Leu Asp Leu Ser Ala Val Pro Ala Ile Asp Thr Ser Gly Ile Ala
141          610          615          620
142  Phe Leu Ile Asp Ile Lys Lys Ser Ile Glu Lys Arg Gly Leu Glu Leu
143          625          630          635          640

```

## RAW SEQUENCE LISTING

DATE: 08/22/2005

PATENT APPLICATION: US/10/762,049

TIME: 12:16:16

Input Set : N:\Crif3\RULE60\10762049.raw

Output Set : N:\CRF4\08222005\J762049.raw

```

144      Val Leu Val Asn Pro Thr Gly Glu Val Met Glu Lys Ile Gln Arg Ala
145                645                650                655
146      Asn Glu Ala Glu Asn Tyr Phe Arg Pro Asp Cys Leu Tyr Leu Thr Thr
147                660                665                670
148      Gly Glu Ala Ile Ala Ser Leu Ser Ala Leu Ala Lys Met Thr Lys Pro
149                675                680                685
151 <210> SEQ ID NO: 3
152 <211> LENGTH: 1981
153 <212> TYPE: DNA
154 <213> ORGANISM: Zea mays
155 <400> SEQUENCE: 3
156      gcacgaggag tccgacctga tcgccggcat caccatcgcc agcctcgcca tcccgagggg 60
157      catcagctac gccaaagctcg ccaacctgcc gcccggtgctc ggactctact cgagcttcgt 120
158      gccgccgctg gtgtacgcgc tgatggggag ctccaaggac ctggcggtgg ggacgggtggc 180
159      ggtggcgctg ctgctcatca gctccatgct cggcagcgag gtgtcgccga cggagaaccc 240
160      cgtgctctac ctgcacctcg ccttcaccgc cactttcttc gccggcgctc tccaggcctc 300
161      gctcggcctc ctgaggttgg gcttcatcgt ggacctgctg tcgcacgcca cgatcgtggg 360
162      gttcatggcc ggcgcgcgga cgggtggtgtg cctgcagcag ctgaaggcca tgctgggcct 420
163      cgtccacttc accacctcca ccgacgtcgt ctccgtcatg gaatccgtct tcagccagac 480
164      acaccagtgg cgggtgggaga gcgtcctgct cgggtgcggc ttctctctct tctctctcgt 540
165      caccgccttc atcagcaaga ggcgtcccaa gctgttctgg atctccgcgg cggcgccggt 600
166      gacgtccgtc gtgctcgga gcgttctggt gtacctcacg cacgtgaaa accacggcat 660
167      cgaagtgatc ggttacctga agaaaggcct gaatccaccg tcggtgacaa gcctgcaatt 720
168      ctcaccgccc tacatgatgc tcgcgtcaa gactgggatc atcaccggcg tcattgccct 780
169      cgccgaagga atcgccgtgg ggaggagctt cgccatgttc aagaactacc acatgacgga 840
170      caacaaggag atgatcgga tcgggacgat gaacgtcctg ggctcgctca cgtcgtgcta 900
171      cctgaccacg gggcccttct cgcgctccgc cgtgaactac aacgccgggt gcaggacggc 960
172      catgtcgaac gtggtcatgt cgctggcggt gatggtcacg ctgctgttcc tgacgccgct 1020
173      gttccactac acgccgctgg tgggtgctgc ggcgatcatc gtctccgcca tgctgggcct 1080
174      ggtcgacttc ggggcgcgc tgcacctgtg gcgcgtcgac aaggctgact tctgcgtctg 1140
175      cgccggcgcg tacctgggcg tcgtcttcgg cagcgtcgag gtcggcctgg tcgtcgccgt 1200
176      cgccgtctcc ctgctccgcg tcctgctgtt cgtcgcccgg cccaggacca cgggtgctcg 1260
177      caacatcccc ggcaccatgg tgtaccggag gatggaccag tacgccgcgg cgcagacggt 1320
178      gccggcgctg ctgctgctgc gcgtcgacgc gcccgctctac ttgcccaacg ccagctacct 1380
179      gcgagagagg atctcgcggt ggatcgacga cgaggaggag cgcaccaaga gccaggcgga 1440
180      gatggcgctg cggtagcttg tctctgacat ggggtgccatc ggtagcatcg acacgagcgg 1500
181      gacgagcatg ctggacgagc tcaacaagtc cttggacagg aggggaatgc agatcgtgct 1560
182      ggcgaacccg ggcagcgaga tcatgaagaa gctggacagc tccaaggtgc tggagcagat 1620
183      cggccacgag tgggtgttcc cgacgggtgg cgaggcggtg gcgtcgtgcg actacgtgct 1680
184      gcactcgcac aagccgggaa tggccaagga cagcgccgcc gccacgaga gcatggtgtg 1740
185      acgagcaccg ccacgccaac cgtatgtgta gtgtgctccg gttccggtct gacgtaacca 1800
186      gtcgtcacgc ggaccgagat gaattatgta tacacgtgtc tcgagtattg tacacctgca 1860
187      ccgtcgcggg aaaaacgaat tcagagaaga aaggatccca cccggttttt tttggtgaaa 1920
188      aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1980
189      a 1981
191 <210> SEQ ID NO: 4
192 <211> LENGTH: 579
193 <212> TYPE: PRT
194 <213> ORGANISM: Zea mays

```

## RAW SEQUENCE LISTING

DATE: 08/22/2005

PATENT APPLICATION: US/10/762,049

TIME: 12:16:16

Input Set : N:\Crif3\RULE60\10762049.raw

Output Set: N:\CRF4\08222005\J762049.raw

195 &lt;400&gt; SEQUENCE: 4

```

196   His Glu Glu Ser Asp Leu Ile Ala Gly Ile Thr Ile Ala Ser Leu Ala
197       1           5           10           15
198   Ile Pro Gln Gly Ile Ser Tyr Ala Lys Leu Ala Asn Leu Pro Pro Val
199           20           25           30
200   Leu Gly Leu Tyr Ser Ser Phe Val Pro Pro Leu Val Tyr Ala Leu Met
201           35           40           45
202   Gly Ser Ser Lys Asp Leu Ala Val Gly Thr Val Ala Val Ala Ser Leu
203           50           55           60
204   Leu Ile Ser Ser Met Leu Gly Ser Glu Val Ser Pro Thr Glu Asn Pro
205           65           70           75           80
206   Val Leu Tyr Leu His Leu Ala Phe Thr Ala Thr Phe Phe Ala Gly Val
207           85           90           95
208   Phe Gln Ala Ser Leu Gly Leu Leu Arg Leu Gly Phe Ile Val Asp Leu
209           100          105          110
210   Leu Ser His Ala Thr Ile Val Gly Phe Met Ala Gly Ala Ala Thr Val
211           115          120          125
212   Val Cys Leu Gln Gln Leu Lys Gly Met Leu Gly Leu Val His Phe Thr
213           130          135          140
214   Thr Ser Thr Asp Val Val Ser Val Met Glu Ser Val Phe Ser Gln Thr
215           145          150          155          160
216   His Gln Trp Arg Trp Glu Ser Val Leu Leu Gly Cys Gly Phe Leu Phe
217           165          170          175
218   Phe Leu Leu Val Thr Arg Phe Ile Ser Lys Arg Arg Pro Lys Leu Phe
219           180          185          190
220   Trp Ile Ser Ala Ala Ala Pro Leu Thr Ser Val Val Leu Gly Ser Val
221           195          200          205
222   Leu Val Tyr Leu Thr His Ala Glu Asn His Gly Ile Glu Val Ile Gly
223           210          215          220
224   Tyr Leu Lys Lys Gly Leu Asn Pro Pro Ser Val Thr Ser Leu Gln Phe
225           225          230          235          240
226   Ser Pro Pro Tyr Met Met Leu Ala Leu Lys Thr Gly Ile Ile Thr Gly
227           245          250          255
228   Val Ile Ala Leu Ala Glu Gly Ile Ala Val Gly Arg Ser Phe Ala Met
229           260          265          270
230   Phe Lys Asn Tyr His Met Thr Asp Asn Lys Glu Met Ile Ala Ile Gly
231           275          280          285
232   Thr Met Asn Val Leu Gly Ser Leu Thr Ser Cys Tyr Leu Thr Thr Gly
233           290          295          300
234   Pro Phe Ser Arg Ser Ala Val Asn Tyr Asn Ala Gly Cys Arg Thr Ala
235           305          310          315          320
236   Met Ser Asn Val Val Met Ser Leu Ala Val Met Val Thr Leu Leu Phe
237           325          330          335
238   Leu Thr Pro Leu Phe His Tyr Thr Pro Leu Val Val Leu Ser Ala Ile
239           340          345          350
240   Ile Val Ser Ala Met Leu Gly Leu Val Asp Phe Gly Ala Ala Leu His
241           355          360          365
242   Leu Trp Arg Val Asp Lys Val Asp Phe Cys Val Cys Ala Gly Ala Tyr
243           370          375          380

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/10/762,049

DATE: 08/22/2005  
TIME: 12:16:17

Input Set : N:\Crf3\RULE60\10762049.raw  
Output Set: N:\CRF4\08222005\J762049.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:9; N Pos. 398,483  
Seq#:10; Xaa Pos. 17,131  
Seq#:14; Xaa Pos. 74

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/762,049

DATE: 08/22/2005

TIME: 12:16:17

Input Set : N:\Crf3\RULE60\10762049.raw

Output Set: N:\CRF4\08222005\J762049.raw

L:417 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9 after pos.:360  
M:341 Repeated in SeqNo=9  
L:436 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10 after pos.:16  
M:341 Repeated in SeqNo=10  
L:521 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14 after pos.:64